

A multifractal approach towards inference in finance

Ola Løvsletten, Martin Rypdal

Department of Mathematics and Statistics, University of Tromsø, N-9037 Tromsø, Norway.

Abstract

We introduce tools for inference in the multifractal random walk introduced by Bacry et al. (2001). These tools include formulas for smoothing, filtering and volatility forecasting. In addition, we present methods for computing conditional densities for one- and multi-step returns. The inference techniques presented in this paper, including maximum likelihood estimation, are applied to data from the Oslo Stock Exchange, and it is observed that the volatility forecasts based on the multifractal random walk have a much richer structure than the forecasts obtained from a basic stochastic volatility model.

Keywords: Multifractal, inference, volatility forecasting, Laplace approximation

1. Introduction

Modeling financial time series by stochastic processes dates back to the work of Bachelier (1900). Bachelier proposed to model the price of a financial asset as a Brownian motion with drift. It was later realized, by e.g. Mitchell (1915), that the standard deviation of price changes are proportional to the price levels themselves. Therefore, Bachelier's model should be modified so that it is the *logarithmic* asset price, $X(t) = \log P(t)$, that is modeled as a Brownian motion with drift. As a modification of this model, Mandelbrot (1963) proposed to replace Brownian motion with α -stable Lévy processes with $\alpha < 2$, so-called Lévy flights.

Both Brownian motions and Lévy flights are selfsimilar and have independent increments. However, empirical analyses of asset prices have revealed that, even though logarithmic returns are uncorrelated, they are nevertheless strongly dependent. This stylized fact is called volatility clustering, and it is not well described by Brownian motions nor Lévy flights. Other processes, such as stochastic volatility (SV) models, are specifically designed to include this feature. The simplest example is the basic SV model of Taylor (1982). If we choose¹ $\mu = 0$, this model is defined by the stochastic differential equation

$$dX(t) = \sigma(t) dB(t),$$

where the logarithmic volatility varies according to an Ornstein-Uhlenbeck process, i.e.

$$d \log \sigma(t) = -a \log \sigma(t) dt + \nu d\tilde{B}(t), \quad (1)$$

where $\tilde{B}(t)$ is a Brownian motion independent of $B(t)$.

Another class of SV models are the multifractal random processes. These models come from turbulence theory, and their origin can be traced back to works of Kolmogorov (1962) and Obukhov (1962). The defining properties of a multifractal process $X(t)$ are stationary increments and structure functions that are power-laws in time, i.e.

$$\mathbb{E}[|X(t)|^q] \sim t^{\zeta(q)}. \quad (2)$$

The scaling functions $\zeta(q)$ are linear for selfsimilar processes, but usually the term “multifractal” refers to the cases where $\zeta(q)$ are strictly concave. For such processes, the absolute values of the increments of $X(t)$ may have algebraically decaying auto-correlation functions (ACFs), even though the increments themselves are uncorrelated. In contrast, the ACFs for the absolute values of the increments decay exponentially in the basic SV models.

That multifractals represent a suitable framework for modeling financial time series was first discovered about fifteen years ago by Ghashghaie et al. (1996) and Mandelbrot et al. (1997). Shortly after this Calvet and Fisher (2001) showed how one can obtain a discrete-time SV model as a discretization of a continuous-time multifractal.

The model constructed by Calvet and Fisher is called the Markov-Switching Multifractal (MSM), and it is con-

Email address: ola.lovsletten@uit.no (Ola Løvsletten)

¹Since μ is easily estimated from data, this can be assumed without any loss of generality.

structed by randomizing the so-called multiplicative cascade. The result is a model that describes log-returns as

$$x_t = \sigma \sqrt{M_t} \varepsilon_t. \quad (3)$$

Here $\varepsilon_t \stackrel{d}{\sim} \mathcal{N}(0, 1)$ are independent variables and the volatility is a product on the form

$$M_t = M_{t,1} M_{t,2} \cdots M_{t,K}.$$

The variables $M_{t,k}$ are updated with different frequencies for different levels k . To be precise, at each time step t , $M_{t,k}$ is given a new value (independently drawn from a distribution M) with probability γ_k and left unchanged with probability $1 - \gamma_k$. The approximate multifractality in the MSM model is achieved by choosing $\gamma_k = 1 - (1 - \gamma_1)^{b^{k-1}}$ for some $\gamma_1 \in (0, 1)$ and some $b > 0$. By exploiting general techniques for Markov-Switching models, Calvet and Fisher have developed inference methods for the MSM model, including maximum likelihood (ML) estimation and volatility forecasting. Unfortunately, there are some limitations to the applicability of these methods. One problem is that the likelihood functions only are available when M is discrete, something that leads to rather unnatural parameterizations. Also, in practice, it is only possible to compute the likelihood if the parameter K does not exceed ≈ 10 (Lux, 2008). In effect, this introduces an unwanted exponential cutoff in the volatility dependence, at the time scale b^K .

At the same time that Calvet and Fisher proposed the MSM model, Bacry et al. (2001) presented a different type of multifractal process, the so-called multifractal random walk (MRW). A popular discrete-time approximation to this process is given by equation (3), with

$$M_t = c e^{h_t}, \quad (4)$$

where h_t is a stationary and centered Gaussian process with co-variances

$$\text{Cov}(h_t, h_s) = \lambda^2 \log^+ \frac{\mathcal{T}}{(|t - s| + 1)\Delta t}. \quad (5)$$

The constant c is chosen so that $1/c = \mathbb{E}[e^{h_t}]$. If the step-length Δt is fixed² it is convenient to denote $R = \mathcal{T}/\Delta t$. The model then depends on three parameters: $\theta = (\lambda, \sigma, R)$.

For the purpose of modeling financial time series, an important property of the MRW model is the slow decay of the volatility dependence. Since the innovations ε_t are

independent, the auto-correlation function for the process $|x_t|$ becomes

$$\mathbb{E}[|x_t x_s|] \propto e^{\frac{1}{2} \mathbb{E}[(h_t/2 + h_s/2)^2]} \propto e^{\frac{1}{4} \text{Cov}(h_t, h_s)}, \quad (6)$$

which for $1 \ll s \ll R$ gives the approximate scaling

$$\mathbb{E}[|x_t x_{t+s}|] \sim s^{-\frac{1}{4} \lambda^2}. \quad (7)$$

The parameter λ is called the intermittency parameter, and it also determines the nonlinearity of the scaling function. In fact, the scaling function of the (continuous-time) MRW model is

$$\zeta(q) = \frac{1}{2} \left(1 + \frac{\lambda^2}{2} \right) q - \frac{\lambda^2}{8} q^2.$$

In contrast to the MSM model, which is obtained by randomizing a discrete multiplicative cascade, the MRW model builds on a continuous cascade. In fact, the log-normal MRW model that we consider in this paper is just a special case of a more general class of processes known as infinitely divisible cascades (Muzy and Bacry, 2002). These processes have very desirable theoretical properties, e.g. exact multifractal scaling. From this point of view, the MRW model is preferable over the MSM model, and it is therefore important to develop inference techniques for the MRW model. A step in this direction was taken in (Løvsletten and Rypdal, 2011), where we presented methods for ML estimation. These results were obtained by observing that the processes defined by equations (3) and (5) are very similar to discrete-time versions of the basic SV models. In fact, if we replace the process h_t with an auto-regressive model of order one (an AR(1) process),

$$h_t = \psi h_{t-1} + \sigma_u u_t, \quad (8)$$

where u_t is Gaussian white noise with unit variance, then the process defined by equations (3) and (4) is a basic SV model. Hence we can use existing techniques for basic SV models (Skaug and Yu, 2009; Martino et al., 2011) in combination with general ML methods for Gaussian processes (McLeod et al., 2007) to obtain likelihoods for the MRW model.

While in (Løvsletten and Rypdal, 2011) we focused on parameter estimation, the focus of this paper is primarily conditional forecasts of returns and inference regarding the latent variables h_t . To be more precise, we are interested in estimating the conditional variables $h_s | \{x_t, t \leq T\}$. For $s < T$ this problem is known as smoothing, for $s = T$ it is called filtering, and for $s > T$ it is called forecasting. These techniques are of obvious importance for the applicability of the MRW model in finance.

²The variable t is dimensionless and represents the number of time steps of length Δt .

The paper is structured as follows. In section 2 we review inference techniques for basic SV models, and in section 3 we generalize these results to the MRW model. In section 4 we apply some of these methods to data from the Oslo Stock Exchange, and in section 5 we give some concluding remarks.

2. Inference techniques in the basic SV model

In general, many statistical problems in stochastic modeling, e.g. model selection, parameter estimation and assessment of uncertainty in estimates, can be solved by utilizing the likelihood of the model. Given data $z = (z_1, \dots, z_T)$, the likelihood \mathcal{L} of a random vector $x = (x_1, \dots, x_T)$, with probability density function $p_x(\cdot|\theta)$, is defined as the function

$$\mathcal{L}(\theta|z) = p_x(z|\theta), \quad (9)$$

i.e. one views the the probability density as a function of the parameters θ , with z fixed.

Remark 1. To simplify notations we will drop the subscripts on the densities throughout the rest of the paper. It will be clear from the arguments which densities are considered. We also suppress the dependency of the parameter vector θ in the notation of the densities.

In the basic SV model the likelihoods are difficult to compute directly. By conditioning on the latent field $h = (h_1, \dots, h_T)$, the probability density of x takes the form

$$p(x) = \int_{\mathbb{R}^T} p(x|h)p(h)dh, \quad (10)$$

where the joint density $p(x, h) = p(x|h)p(h)$ is a product of the Gaussian marginals

$$p(x|h) = \prod_{i=1}^T p(x_i|h_i) \quad (11)$$

and

$$p(h) = p(h_1) \prod_{i=2}^T p(h_i|h_{i-1}). \quad (12)$$

The factors in equations (11) and (12) are densities corresponding to the distributions $x_t|h_t \stackrel{d}{\sim} \mathcal{N}(0, \sigma^2 c \exp(h_t))$, $h_1 \stackrel{d}{\sim} \mathcal{N}(0, \sigma_u^2/(1 - \psi^2))$ and $h_t|h_{t-1} \stackrel{d}{\sim} \mathcal{N}(\psi h_{t-1}, \sigma_u^2)$. In general the integral in equation (10) has no closed form, and it is typically very demanding to compute numerically. As an approximation one may consider a second-order Taylor expansion of $\log p(x, h)$ around the maximum

$$h^* = \operatorname{argmax}_h \log p(x, h). \quad (13)$$

The resulting integral is easily computed, giving the expression

$$p_x(x) \approx (2\pi)^{T/2} |\det \Omega(x)|^{-1/2} p(x, h^*), \quad (14)$$

where

$$\Omega(x) = \frac{\partial^2 \log p(x, h)}{\partial h^T \partial h} \Big|_{h=h^*} \quad (15)$$

is the Hessian matrix of the map $h \mapsto \log p(x, h)$, evaluated at $h = h^*$. This approximation is known as Laplace's method, and it has been applied, by among others Martino et al. (2011), to compute likelihood functions in basic SV models. The reason for its efficiency in basic SV models is the Markov structure of the latent field h . The Markov property ensures that the gradient of $h \mapsto \log p(x, h)$ is on the form

$$\frac{\partial \log p(x, h)}{\partial h_i} = b_i + \sum_j A_{ij} h_j + g_i(x_i, h_i), \quad (16)$$

where $A = \|A_{ij}\|$ is a tridiagonal matrix, b_i are constants and g_i are non-linear functions. By exploiting the sparseness of A , one can efficiently calculate h^* . In addition, the Hessian matrix $\Omega(x)$ is tridiagonal, making the computation of the expression in equation (14) efficient.

We are now in a position to make statistical inference based on the basic SV model. We start by looking at filtering of the volatilities. Overlooking model uncertainty and parameter uncertainty, the conditional density $p(h_T|x)$ contains all available information about the latent variable h_T at time T . As a point estimate one may consider the Bayes estimator which is defined as the maximum of the posterior distribution $p(h_T|x)$. This density is approximated using Laplace's method :

$$\begin{aligned} p(h_T|x) &\propto p(x, h_T) \\ &= \int_{\mathbb{R}^{T-1}} p(x, h) dh_1 \cdots dh_{T-1} \\ &\approx b p(x, \tilde{h}_1, \dots, \tilde{h}_{T-1}, h_T), \end{aligned} \quad (17)$$

where the factor b does not depend on h_T , and

$$(\tilde{h}_1, \dots, \tilde{h}_{T-1}) = \operatorname{argmax}_{h_1, \dots, h_{T-1}} \log p(x, h). \quad (18)$$

Maximizing (17) gives the filtered estimate \hat{h}_T of h_T . The filtering procedure can be written more compact as in (13) with $\hat{h}_T = h_T^*$.

For smoothing we consider the posterior distribution $p(h_s|x)$, now with $s < T$. A similar argument as for the derivation of the filtering formula gives the approximated Bayes estimator $\hat{h}_s = h_s^*$ where h_s^* is component s of the vector h^* in equation (13).

We note that if we are already calculating the likelihood using the approach described above, then very little additional effort is required to obtain these estimates, since the maxima in equation (13) is found as a part of the Laplace approximation.

To forecast the volatility N steps into the future we follow the same procedure as for smoothing and filtering. We need to find the maximum of the expression

$$\log p(x, h, h_{T+N}) = \log p(x, h) + \log p(h_{T+N}|h)$$

as a function of (h, h_{T+N}) . Iterating equation (8) backwards yields

$$h_{T+N} = \psi^N h_T + \sum_{k=0}^{N-1} \psi^k u_{T+N-k},$$

and hence

$$h_{T+N}|h \stackrel{d}{\sim} \mathcal{N}\left(\psi^N h_T, \sigma_u^2 \sum_{k=0}^{N-1} \psi^{2k}\right).$$

Differentiation of $\log p(h_{T+N}|h)$ gives

$$\frac{\partial \log p(h_{T+N}|h)}{\partial h_{T+N}} = -\frac{(h_{T+N} - \psi^N h_T)}{\sigma_u^2 \sum_{k=0}^{N-1} \psi^{2k}}. \quad (19)$$

To find a maximum we require that the expressions in equation (19) equals zero, and also that $\nabla_h \log p(x, h) = 0$. From this, the N -step volatility forecast becomes

$$\hat{h}_{T+N} = \psi^N \hat{h}_T, \quad (20)$$

where \hat{h}_T is the filtered estimate of h_T . The formulas we have derived for smoothing, filtering and forecasting of the volatilities are the same as in Skaug and Yu (2009).

To conclude this section we remark that, since $p(x_{T+N}|x) \propto p(x, x_{T+N})$, the conditional densities $p(x_{T+N}|x)$ can be computed simply by using the Laplace approximation. For $N > 1$ one must take into account that the matrices A and $\Omega(x, x_{T+N})$ are modified due to the inclusion of the density $p(h_{T+N}|h)$.

3. Generalization to the MRW model

In this section we extend the results of section 2 to the discrete-time MRW model. In this case h_t is no longer a Markov process. While h_t still is a centered Gaussian process, its covariance structure is now given by equation (5).

Let us first review the approximation of the likelihood for the MRW model (Løvsletten and Rypdal, 2011). One

starts with Laplace's method, given in equation (14). The density $p(x|h)$ is the same as for the basic SV model, but the density of $p(h)$ needs to be handled differently. We denote by $\gamma(k) = \text{Cov}(h_0, h_k)$ the auto-covariance function of the process h_t , and let Γ_t be the variance-covariance matrices of the vectors (h_1, \dots, h_t) . That is

$$\Gamma_t(i, j) = \gamma(|i - j|) \text{ for } i, j = 0, 1, \dots, t-1.$$

As usual when working with Gaussian vectors, it is convenient to introduce regression coefficients $\phi_i^{(t)}$. The vectors $\phi^{(t)}$ are defined via the equations

$$\Gamma_t \phi^{(t)} = \gamma_{1:t}, \quad (21)$$

where $\gamma_{1:t} = (\gamma(1), \dots, \gamma(t))^T$. From standard theory of multivariate normal distributions, the conditional distributions of $h_t | \{h_s : 1 \leq s < t\}$ are normal,

$$h_t | \{h_s : 1 \leq s < t\} \stackrel{d}{\sim} \mathcal{N}(m_t, P_t), \quad (22)$$

with means $m_t = (h_{t-1}, h_{t-2}, \dots, h_1) \phi^{(t-1)}$ and variances $P_t = \gamma(0) - \gamma_{1:t-1}^T \Gamma_{t-1}^{-1} \gamma_{1:t-1}$. Since the density of h can be decomposed into a product of one-dimensional marginals, equation (22) gives $p(h)$.

Remark 2. Solving the the equations in (21) for $t = 1, \dots, T-1$ can be done iteratively using the Durbin-Levinson algorithm, which requires only $O(T^2)$ floating point operations. We refer to (McLeod et al., 2007) for details.

A second difference between the basic SV model and the MRW model is the structure of the matrices Ω and A , which are defined by equations (15) and (16). For the MRW model these are no longer sparse. This makes the computation of the expression in equation (14) extremely demanding. The solution is to truncate the dependency in the process h_t after a finite number of lags. This gives the approximation:

$$p(h_t | \{h_s : 1 \leq s < t\}) \approx p(h_t | \{h_s : t - \tau \leq s < t\}), \quad (23)$$

where $\tau \in \mathbb{N}$ is a truncation parameter. We note that for $t > \tau$, the regression coefficients and variances of $h_t | \{h_s : t - \tau \leq s < t\}$ are $\phi^{(\tau)}$ and $P_{\tau+1}$ respectively. After truncation, the matrices A and Ω become band-diagonal with bandwidths equal τ .

Remark 3. The likelihood approximation for the MRW model is implemented in the R computer language. In our implementation we have used analytical expressions for the first and second order derivatives to construct the matrices Ω and A . The maxima h^* are found by numerically

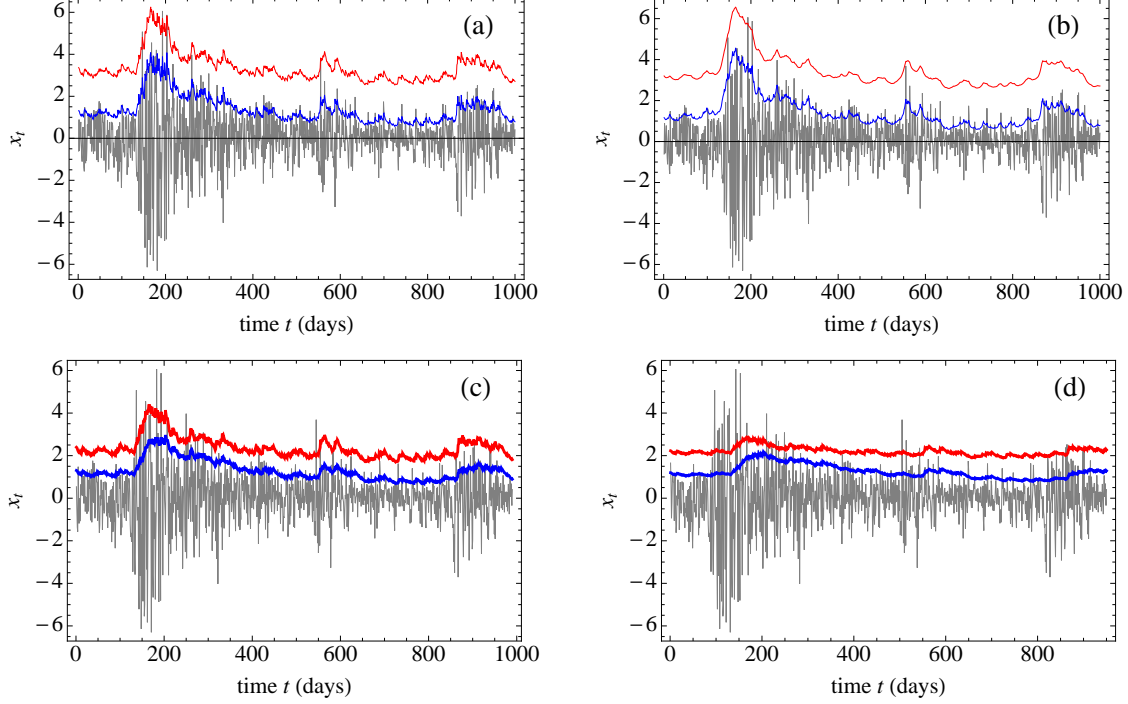


Figure 1: (a): Shows $e^{\hat{h}_t/2}$, where \hat{h}_t are the filtered estimates of h_t from the daily log-returns in the OSEBX in the time period from February 20th 2008 to February 8th 2012. The lower curve is for the MRW model, and the top curve is for the basic SV model. The top curve has been shifted to make it visible. The filtered signals are plotted together with the log-returns x_t of the OSEBX. (b): Shows the same as in (a), but now \hat{h}_t are the smoothed estimates of h_t given all the observations of x_t . (c): Shows $e^{\hat{h}_t/2}$, where \hat{h}_t is the forecast performed using data $\{x_s : s \leq t - N\}$ with $N = 10$ days. (d): Shows the same as in (c), but now with $N = 50$ days.

calculating the roots of the expressions in equation (16) using the algorithm "DF-SANE" (La Cruz et al., 2006). This algorithm is implemented in the R package "BB" (Varadhan and Gilbert, 2009). To find the determinant of the Ω we use the package "Matrix" which efficiently stores and manipulates sparse matrices.

With the likelihood approximation at hand, we can extend the formulas for smoothing, filtering and forecasting to the MRW model. As for the basic SV model, we maximize the posterior distribution according to equation (13), and the formulas for smoothing and filtering are exactly as for the basic SV model.

To forecast the volatilities N steps ahead we need the conditional density of $h_{T+N}|h$. Since this variable is normal, the distribution is uniquely given by the mean $m_{T+N|T}$ and variance $P_{T+N|T}$, i.e.

$$h_{T+N}|h \stackrel{d}{\sim} \mathcal{N}(m_{T+N|T}, P_{T+N|T}). \quad (24)$$

The mean is a linear combination the conditioning variables, i.e.

$$m_{T+N|T} = (h_T, h_{T-1}, \dots, h_1) \phi^{(T,N)},$$

where the coefficients $\phi^{(T,N)}$ are solutions to the equations

$$\Gamma_T \phi^{(T,N)} = \gamma_{N:T+N-1}, \quad (25)$$

with

$$\gamma_{N:T+N-1} = (\gamma(N), \gamma(N+1), \dots, \gamma(N+T-1))^T.$$

The variance is given by

$$P_{T+N|T} = \gamma(0) - \gamma_{N:T+N-1}^T \Gamma_T \gamma_{N:T+N-1}.$$

We note that in the special case $N = 1$ we have $\phi^{(T,N)} = \phi^{(T)}$, and we can again use the Durbin-Levinson algorithm. In the case $N > 1$, the explicit inverse of Γ_T is needed, and one may use the algorithm of Trench (1964), which utilizes that the matrices are Toeplitz. Using the same procedure as in section 2, we get the forecasting formula

$$\hat{h}_{T+N} = (\hat{h}_T, \hat{h}_{T-1}, \dots, \hat{h}_1) \phi^{(T,N)}, \quad (26)$$

where $(\hat{h}_1, \hat{h}_2, \dots, \hat{h}_T)$ are the smoothed estimates of (h_1, h_2, \dots, h_T) .

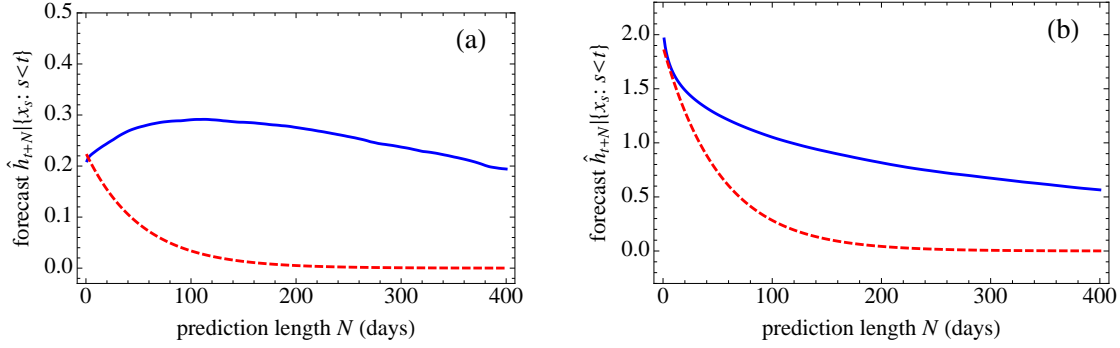


Figure 2: (a): Examples of forecasts \hat{h}_{t+N} for the basic SV model (dotted curve) and the MRW model (solid curve). The forecasts are computed from the OSEBX data with the time t corresponding to the date February 17th 2010. (b): Same as in (a), but now the forecasts are performed for the date March 5th 2009.

Using the Laplace approximation for the MRW model, the N step conditional densities $p(x_{T+N}|x)$ are computed as in section 2.

4. Examples

As an example we have applied the inference methods presented in sections 2 and 3 to a time series consisting of daily log-returns of the Oslo Stock Exchange Benchmark Index (OSEBX). The data used are closing prices for the time period May 25th 2001 to February 8th 2012, and the whole time series is used to obtain ML estimates for the basic SV model and the MRW model. The interesting estimates are $\hat{\psi} = 0.98$ for the basic SV model and $\hat{\lambda} = 0.33$ for the MRW model.

In figure 1(a) we have plotted the filtered estimates of h_t together with the log-returns for the time period from February 20th 2008 to February 8th 2012. The filtering for the basic SV model and the MRW model are similar, but not identical. The same is seen in figure 1(b), which shows the smoothed estimates. In figures 1(c) and 1(d) we have plotted the N -step forecast for $N = 10$ days and $N = 50$ days respectively. In the 50-day forecast there are some clear visible differences between the two models. These differences become even clearer in figure 2. In this figure we show two examples, where we (for a fixed time t) make future predictions \hat{h}_{t+N} , and plot these as a functions of N . It follows from equation (20) that these curves must be monotonic and exponentially decaying for the basic SV model. This is not the case for the MRW model, and we observe that the forecasts based on the this model have much richer behavior. We note that similar observations

have been made for the MSM model (Calvet and Fisher, 2001).

As explained in sections 2 and 3, it is possible to use the Laplace approximation to compute the full conditional densities for future returns. This gives forecasts containing more information than the estimates presented in figures 1 and 2. In figure 3 we show two examples where such densities have been computed. In these examples, the volatility is high, and the conditional densities are wider than the unconditioned density. In other situations, where the volatility is low, the conditional densities will be narrower than the unconditioned density.

Remark 4. The computer code that is used for these examples is available online at complexityandplasmas.net.

5. Conclusion

The main results of this paper are methods for smoothing, filtering and forecasting using the MRW model. In addition, we have presented methods for computing conditional densities of future returns. These results improve on existing forecasting techniques for multifractal models, and we therefore consider this work to be an important contribution to the field.

The methods presented in this work open the way for several future studies of multifractal modeling in finance. Among the new possibilities that we consider most interesting, are model comparisons based on estimated future distributions.

Acknowledgement

This project is supported by *Sparebank 1 Nord-Norge*.

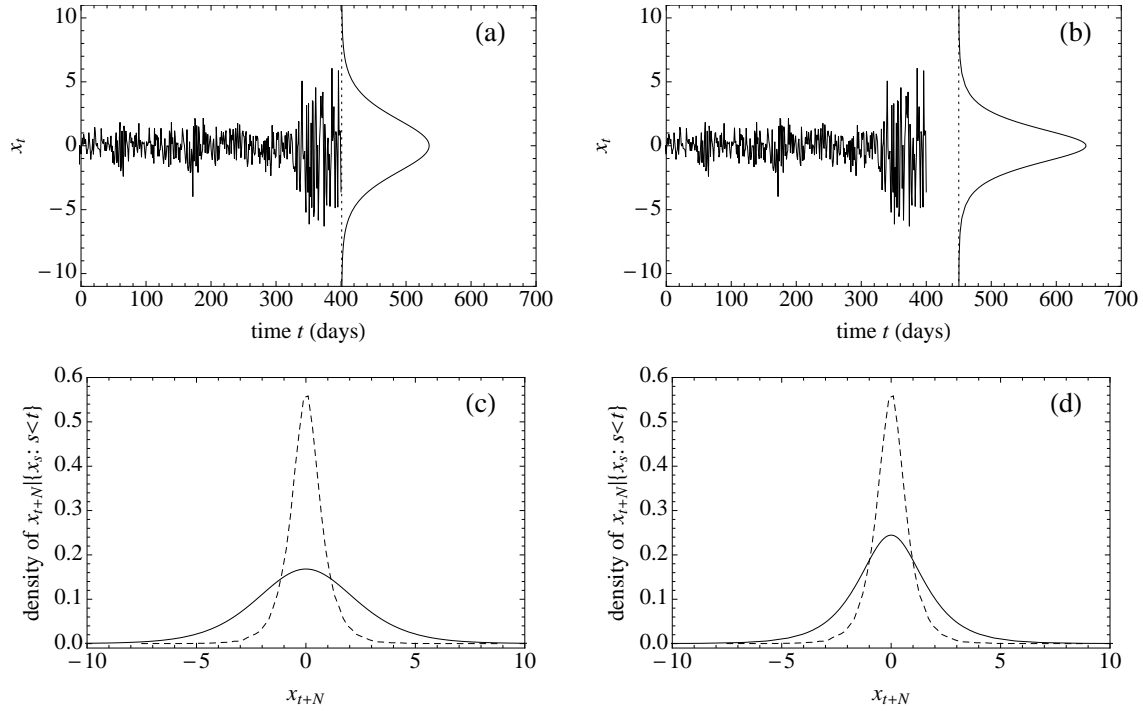


Figure 3: (a): An illustration of the use of conditional densities in forecasting. The figure shows the log-returns of the OSEBX up to December 12th 2008 and the conditional density of the log-return for the next day. (b): Shows the same as (a), but now the conditional density is a N -step forecast with $N = 50$ days. (c): Shows the same conditional density (solid curve) as is illustrated in (a). This density is compared with the unconditional density for the log-returns (dashed line). (d): Same as in (c), but now for the N -step forecast with $N = 50$ days

References

References

- Bachelier, L., 1900. Théorie de la spéculation. Annales scientifiques de l'É.N.S 17, 21–86.
- Bacry, E., Delour, J., Muzy, J. F., 2001. Multifractal random walk. Physical Review E 64, 026103.
- Calvet, L., Fisher, A., Nov. 2001. Forecasting multifractal volatility. Journal of Econometrics 105 (1), 27–58.
- Ghashghaie, S., Breymann, W., Peinke, J., Talkner, P., Dodge, Y., 1996. Turbulent Cascades in Foreign-Exchange Markets. NATURE 381, 767–770.
- Kolmogorov, A. N., 1962. A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. Journal of Fluid Mechanics 13, 82–85.
- La Cruz, W., Martínez, J. M., Raydan, M., 2006. Spectral residual method without gradient information for solving large-scale nonlinear systems of equations. Mathematics of Computation 75, 1429–1448.
- Løvsletten, O., Rypdal, M., 2011. Approximated maximum likelihood estimation in multifractal random walks. arXiv.org physics.data-an.
- Lux, T., 2008. The Markov-Switching Multifractal Model of Asset Returns: GMM Estimation and Linear Forecasting of Volatility. Journal of Business & Economic Statistics 26, 194–210.
- Mandelbrot, B., 1963. The Variation of Certain Speculative Prices. The Journal of Business 36, 394.
- Mandelbrot, B., Fisher, A., Calvet, L., 1997. A Multifractal Model of Asset Returns. Cowles Foundation Discussion Paper 1164.
- Martino, S., Aas, K., Lindqvist, O., Neef, L. R., Rue, H., 2011. Estimating stochastic volatility models using integrated nested Laplace approximations. The European Journal of Finance 17 (7), 487–503.
- McLeod, I. A., Yu, H., Krougly, Z. L., 2007. Algorithms for Linear Time Series Analysis: With R Package . Journal of Statistical Software 23 (5).

- Mitchell, W. C., 1915. The making and using of index numbers. *Bulletin of the United States Bureau of Labor Statistics* 173, 5–114.
- Muzy, J. F., Bacry, E., 2002. Multifractal stationary random measures and multifractal random walks with log infinitely divisible scaling laws. *Physical Review E* 66, 056121.
- Obukhov, A. M., 1962. Some Specific Features of Atmospheric Turbulence. *Journal of Geophysical Research* 67, 3011–3014.
- Skaug, H. J., Yu, J., 2009. Automated Likelihood Based Inference for Stochastic Volatility Models. Singapore Management University, School of Economics Working Paper, 1–25.
- Taylor, S. J., 1982. Financial Returns Modelled by the Product of Two Stochastic Processes—A Study of Daily Sugar Prices, 1961–79. In: *Time Series Analysis: Theory and Practice*. North Holland, pp. 203–226.
- Trench, W. F., 1964. An Algorithm for the Inversion of Finite Toeplitz Matrices. *Journal of the Society for Industrial and Applied Mathematics* 12, 515–522.
- Varadhan, R., Gilbert, P., 2009. BB: An R Package for Solving a Large System of Nonlinear Equations and for Optimizing a High-Dimensional Nonlinear Objective Function. *Journal of Statistical Software* 32, 1–26.